Combined uncertainty factor for sampling and analysis

Article (Published Version)

Accred Qual Assur
DOI 10.1007/s00769-017-1271-y

GENERAL PAPER

Combined uncertainty factor for sampling and analysis

Michael H. Ramsey[1] · **Stephen L. R. Ellison**[2]

Received: 17 March 2017 / Accepted: 16 May 2017

Abstract Measurement uncertainty that arises from primary sampling can be expressed as an uncertainty factor, which recognises its sometimes approximately log-normal probability distribution. By contrast, uncertainty arising from chemical analysis is usually expressed as relative uncertainty, based upon the assumptions of its approximately normal distribution. A new method is proposed that enables uncertainty from these two sources, expressed in these different ways, to be combined to produce an estimate of the total combined uncertainty of the measurement values that result when the measurement process is considered as a whole.

Keywords Measurement uncertainty · Uncertainty factor · Sampling uncertainty · Propagation of uncertainty

Introduction

A case has recently been made for expressing measurement uncertainty as an uncertainty factor (u' or UF) in place of relative uncertainty (as a percentage) in cases where the relative standard uncertainty exceeds 20 % to 30 % [1]. This is because the frequency distribution of the uncertainty under these conditions has often been found to be roughly log-normal. The dominant source of such high levels of uncertainty often arises from the process of primary sampling rather than from the chemical analysis. However, a potential problem arises in calculating the combined measurement uncertainty when the probability distributions of the two main sources are different, with the analytical source being close to normal and the sampling source being quasi-log-normal. This document proposes a novel procedure by which this calculation can be made, and hence a combined uncertainty factor estimated.

✉ Michael H. Ramsey
m.h.ramsey@sussex.ac.uk

[1] University of Sussex, Brighton, UK

[2] LGC, Teddington, UK

Method

Usually the combination of the standard measurement uncertainty (u_{meas}) arising from chemical analysis (u_{anal}) and that from sampling (u_{samp}) has been effected by adding the estimated variances s^2_{anal} and s^2_{samp}; thus,

$$u_{\text{meas}} = s_{\text{meas}} = \sqrt{s^2_{\text{anal}} + s^2_{\text{samp}}}\,.$$

This assumes that no uncertainty arises from systematic effects, such as bias in the sampling process. A comparable type of relationship can also be applied to the uncertainties expressed as relative standard deviations ($s_{\text{rel,meas}} = s_{\text{meas}}/\bar{x}$), in relation to the arithmetic mean ($\bar{x}$)

$$s_{\text{rel,meas}} = \sqrt{s^2_{\text{rel,anal}} + s^2_{\text{rel,samp}}}\,.$$

Studies in sampling have shown that at relatively high levels of uncertainty the frequency distribution of results is better described by a log-normal, rather than normal, distribution [2], in the sense that the results do not contain negative or zero values, and usually show a positive skew. When a normal distribution is assumed, and the uncertainty is large (e.g. over 50 % as a relative standard uncertainty), it is possible to have confidence limits of the measurand that include negative concentration values that are clearly impossible. Measurement uncertainty, such as that from

Published online: 26 May 2017

sampling, can therefore be more usefully quantified by using the standard deviation (s) of the natural logarithms of the measured concentration values (x)

$$s_G = s(\ln(x)).$$

This is closely related to the geometric standard deviation (s_g), which is defined as the antilog of s_G, and can be expressed as

$$s_g = \exp(s(\ln(x))).$$

For the summation of uncertainties in log-space, if both components of the uncertainty are expressed in the form of s_G, they can be summed using

$$s_{G,meas} = \sqrt{s_{G,samp}^2 + s_{G,anal}^2}. \quad (1)$$

The geometric standard deviation (s_g) has also been proposed as an uncertainty factor (u') by which the measurement value can be either multiplied, or divided, to calculate the upper and lower confidence limits, respectively, in linear space [1]. The uncertainty factor of the measurement is therefore given by

$$u'_{meas} = \exp\left(s_{G,meas}\right) = \exp\sqrt{s_{G,samp}^2 + s_{G,anal}^2}. \quad (2)$$

This uncertainty factor can also be calculated by summing those of its components using the more complex expression

$$u'_{meas} = \exp\sqrt{(\ln(u'_{samp}))^2 + (\ln(u'_{anal}))^2}$$

which, in terms of geometric standard deviations, is equivalent to

$$s_{g,meas} = \exp\sqrt{\left(\ln\left(s_{g,samp}\right)\right)^2 + \left(\ln\left(s_{g,anal}\right)\right)^2}.$$

However, it is widely observed that analytical uncertainty is usually distributed in a form that approximates to the Gaussian or normal distribution [3]. It is proposed that the summation of these uncertainties, when they are expressed in different forms, can be approached by using an approximation when $s_{rel,anal} < 0.2$, namely

$$s_{G,anal} \approx s_{rel,anal}. \quad (3)$$

The case for this approximation is given in the GUM [4], where we have

$$s(\ln(x)) \approx \left[\frac{d(\ln(x))}{dx}\right] s(x) = \left(\frac{1}{x}\right) s(x) = s_{rel}(x)$$

and hence for small $s(x)$

$$s(\ln(x)) = s_G \approx s_{rel}(x).$$

This gives the approximation to Eq. (1), in which

$$s_{G,meas} \approx \sqrt{s_{G,samp}^2 + s_{rel,anal}^2}. \quad (4)$$

For the complete measurement process from primary sampling to analytical determination, estimates of the measurement uncertainty can be made by the 'duplicate method', followed by analysis of variance [5]. When both the analytical and sampling uncertainties are assumed to be log-normally distributed, the ANOVA can be applied to the natural logarithms of the raw concentration values, prior to use of Eq. (2). When estimates of analytical uncertainty have been made, assuming it has a normal distribution, the values of s_{rel} values, calculated from the raw concentration values if $s_{rel,anal} < 0.2$, are effectively identical to those of s_G, from Eq. (3). The summation of uncertainties expressed in these two different forms can then be made using a modified version of Eq. (2)

$$u'_{meas} = \exp\sqrt{s_{G,samp}^2 + s_{rel,anal}^2}. \quad (5)$$

An example of the application of this approach uses the measurements of lead concentration in very heterogeneous soil at a 9-ha site, as part of the assessment of land for potential housing development [5]. The measurement uncertainty was estimated using the 'duplicate method' followed by classical ANOVA on the natural logarithms of the concentration values for ten duplicate samples. The values of s_G for sampling and analysis were 0.4784 and 0.0567, respectively. This gives the value of $s_{G,meas}$ as 0.4817 [using Eq. (1)] and u'_{meas} as 1.6189 [using Eq. (2)]. Alternatively, considering the analytical uncertainty as normally distributed, it has a $s_{rel,anal}$ value of 0.0566 (calculated from classical ANOVA of the raw measurement values), which gives a virtually identical value for u'_{meas}, (1.6188) using Eq. (5). A typical measured lead concentration value of 300 mg/kg, for example, would therefore have an asymmetric 95 % confidence interval of 93 mg/kg to 971 mg/kg [i.e. $300/(1.619 \times 2)$ to $300 \times (1.619 \times 2)$].

This new approach will enable more reliable estimates of measurement uncertainty, and the consequent confidence limits, especially when the level of sampling uncertainty is high and log-normally distributed. The issue of which of the two frequency distributions to assume for the analytical uncertainty makes little difference to the overall measurement uncertainty (assuming $s_{rel,anal} < 0.2$, as is usually the case). However, the use of the log-normal distribution is probably more reliable as a default option, as it does not rely on this assumption.

References

1. Ramsey MH, Ellison SLR (2015) Uncertainty factor: an alternative way to express measurement uncertainty in chemical measurement. Accred Qual Assur 20:153–155. doi:10.1007/s00769-015-1115-6
2. Whitaker TB, Dowell FE, Hagler WM, Giesbrecht FG, Wu J (1994) Variability associated with sampling, sample preparation and chemical testing for aflatoxin in farmer's stock peanuts. J AOAC Int 77(1):107–116
3. Thompson M, Howarth RJ (1980) The frequency distribution of analytical error. Analyst 105:1188–1195
4. JCGM 101:2008 (2008) Evaluation of measurement data—supplement 1 to the "Guide to the expression of uncertainty in measurement"—propagation of distributions using a Monte Carlo method. http://www.bipm.org/utils/common/documents/jcgm/JCGM_101_2008_E.pdf
5. Ramsey MH, Ellison SLR (eds) (2007) Eurachem/EUROLAB/CITAC/Nordtest/AMC guide: measurement uncertainty arising from sampling: a guide to methods and approaches. Eurachem ISBN:978 0 948926 26 6